On Channel Estimation Using Superimposed Training and First-Order Statistics

Jitendra K. Tugnait, *Fellow, IEEE*, and Weilin Luo

***Abstract*—Channel estimation for single-input multiple-output (SIMO) time-invariant channels is considered using only the first-order statistics of the data. A periodic (nonrandom) training sequence is added (superimposed) at a low power to the information sequence at the transmitter before modulation and transmission. Recently superimposed training has been used for channel estimation assuming no mean-value uncertainty at the receiver and using periodically inserted pilot symbols. We propose a different method that allows more general training sequences and explicitly exploits the underlying cyclostationary nature of the periodic training sequences. We also allow mean-value uncertainty at the receiver. Illustrative computer simulation examples are presented.**

***Index Terms*—Channel estimation, superimposed training.**

I. Introduction

CONSIDER an single-input multiple-output (SIMO) finite-impulse response (FIR) linear channel with N outputs. Let $\{s(n)\}$ denote a scalar sequence which is input to the SIMO channel with discrete-time impulse response $\{\mathbf{h}(l)\}$. The vector channel may be the result of multiple receive antennas and/or oversampling at the receiver. Then the symbol-rate, channel output vector is given by

$$\mathbf{x}(n) := \sum_{l=0}^{L} \mathbf{h}(l)s(n-l). \tag{1}$$

The noisy measurements of $\mathbf{x}(n)$ are given by

$$\mathbf{y}(n) = \mathbf{x}(n) + \mathbf{v}(n). \tag{2}$$

A main objective in communications is to recover $s(n)$ given noisy $\{\mathbf{x}(n)\}$. In several approaches this requires knowledge of the channel impulse response [3], [5]. In training-based approach, $s(n) = c(n) =$ training sequence (known to the receiver) for (say) $n = 1, 2, \ldots, M$ and $s(n)$ for $n > M$ is the information sequence (unknown apriori to the receiver) [3], [5]. Therefore, given $c(n)$ and corresponding noisy $\mathbf{x}(n)$, one estimates the channel via least-squares and related approaches. For time-varying channels, one has to send training signal frequently and periodically to keep up with the changing channel. This wastes resources. An alternative is to estimate the channel based solely on noisy $\mathbf{x}(n)$ exploiting statistical and other properties of $\{s(n)\}$[3], [5]. This is the blind channel estimation approach. More recently, [1] and [2] have explored a superimposed training based approach for time-invariant systems where one takes $s(n) = c(n) + b(n)$, $\{b(n)\}$ is the information sequence and $\{c(n)\}$ is a nonrandom periodic training (pilot) sequence. Exploitation of the periodicity of $\{c(n)\}$ allows identification of the channel without allocating any explicit time slots for training, unlike traditional training methods. There is no loss in information rate. On the other hand, some useful power is wasted in superimposed training which could have otherwise been allocated to the information sequence. This lowers the effective signal-to-noise ratio (SNR) for the information sequence and affects the bit error rate (BER) at the receiver.

Let

$$s(n) = b(n) + c(n) \tag{3}$$

in (1) where $\{b(n)\}$ is the information sequence and $c(n)$ is the superimposed training sequence. Let $\delta(\tau)$ denote the Kronecker delta, I_N denote the $N \times N$ identity matrix and the superscript H denote the complex conjugate transpose operation. Assume the following:

(H1) the information sequence $\{b(n)\}$ is zero-mean, white with $E\{|b(n)|^2\} = 1$;

(H2) the measurement noise $\{\mathbf{v}(n)\}$ is **nonzero-mean** ($E\{\mathbf{v}(n)\} = \mathbf{m}$), white, uncorrelated with $\{b(n)\}$, with $E\{[\mathbf{v}(n+\tau) - \mathbf{m}][\mathbf{v}(n) - \mathbf{m}]^H\} = \sigma_v^2 I_N \delta(\tau)$. The mean vector $\mathbf{m}$ is unknown;

(H3) the superimposed training sequence $c(n) = c(n + mP)\ \forall m, n$ is a nonrandom periodic sequence with period P.

Reference [1] uses the second-order statistics of the received signal to estimate the channel whereas [2] exploits the first-order statistics. As in [2] we will exploit the first-order statistics of the received signal. (A consequence of using the first-order statistics is that the knowledge of the noise variance σ_v^2 in (**H2**) is not used.) The corresponding time-invariant model in [2] (also [1]) does not include an unknown constant term (d.c. offset) in the measurement equation [$\mathbf{m}$ in (**H2**)]; it should, however, if we exploit $E\{\mathbf{y}(n)\}$ to estimate the channel. In practice, linear systems arise because of linearization about some operating (set) point–"bias" in amplifiers, e.g., These set points are typically unknown (at least not known precisely) *a priori*, and one does not normally worry about them since unknown means are estimated and removed before processing (blocked by capacitor-coupling etc.) and they are not needed in any processing. However, if (time-varying) mean $E\{\mathbf{y}(n)\}$ is what we wish to use (as in [2]), then we must include a term such as nonzero $\mathbf{m}$. Reference [2] proposes the choice $c(n) = \sum_k a\delta(n - kP)$. The choice of [2] leads to a poor peak-to-average power ratio of the

Manuscript received February 26, 2003. The associate editor coordinating the review of this letter and approving it for publication was Prof. P. Loubaton. This work was supported by the Army Research Office under Grant DAAD19-01-1-0539.

The authors are with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: tugnait@eng.auburn.edu; wlluo@eng.auburn.edu).

Digital Object Identifier 10.1109/LCOMM.2003.817325

transmitted signal which is highly undesirable if the transmit power amplifier has some nonlinearity. In this paper we follow the basic ideas of [1] and [2] but propose a different method which works for nonzero $\mathbf{m}$ in (**H2**).

II. Superimposed Training-Based Solution

By (1)–(2) and (**H3**), we have

$$E\{\mathbf{y}(n)\} = E\{\mathbf{x}(n)\} + \mathbf{m} = \sum_{l=0}^{L} \mathbf{h}(l)c(n-l) + \mathbf{m}. \quad (4)$$

Since $\{c(n)\}$ is periodic, we have ($\alpha_m := 2\pi m/P$)

$$c(n) = \sum_{m=0}^{P-1} c_m e^{j\alpha_m n} \;\; \forall n, \;\; c_m := \frac{1}{P}\sum_{n=0}^{P-1} c(n)e^{-j\alpha_m n}. \quad (5)$$

The coefficients c_m's are known at the receiver since $\{c(n)\}$ is known. We have

$$E\{\mathbf{y}(n)\} = \sum_{m=0}^{P-1} \underbrace{\left[\sum_{l=0}^{L} c_m \mathbf{h}(l) e^{-j\alpha_m l}\right]}_{=:\mathbf{d}_m} e^{j\alpha_m n} + \mathbf{m}. \quad (6)$$

The sequence $E\{\mathbf{y}(n)\}$ is periodic [4] with cycle frequencies α_m, $0 \le m \le P-1$. A mean-square (m.s.) consistent estimate $\hat{\mathbf{d}}_m$ of $\mathbf{d}_m$, for $\alpha_m \ne 0$, follows as [5]

$$\hat{\mathbf{d}}_m = \frac{1}{T}\sum_{n=1}^{T} \mathbf{y}(n)e^{-j\alpha_m n}. \quad (7)$$

As $T \to \infty$, $\hat{\mathbf{d}}_m \to \mathbf{d}_m$ m.s. if $\alpha_m \ne 0$ and $\hat{\mathbf{d}}_0 \to \mathbf{d}_0 + \mathbf{m}$ m.s. if $\alpha_m = 0$.

We now establish that given $\mathbf{d}_m$ for $1 \le m \le P-1$, we can (uniquely) estimate $\mathbf{h}(l)$'s if $P \ge L+2$, $\alpha_m \ne 0$, and $c_m \ne 0 \; \forall m \ne 0$. Since $\mathbf{m}$ is unknown, we will omit the term $m = 0$ for further discussion. Define

$$\mathbf{V} := \begin{bmatrix} 1 & e^{-j\alpha_1} & \cdots & e^{-j\alpha_1 L} \\ 1 & e^{-j\alpha_2} & \cdots & e^{-j\alpha_2 L} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\alpha_{P-1}} & \cdots & e^{-j\alpha_{P-1} L} \end{bmatrix}_{(P-1)\times(L+1)} \quad (8)$$

$$\mathcal{H} := [\mathbf{h}^H(0) \;\; \mathbf{h}^H(1) \;\; \cdots \;\; \mathbf{h}^H(L)]^H \quad (9)$$

$$\mathcal{D} := [\mathbf{d}_1^H \;\; \mathbf{d}_2^H \;\; \cdots \;\; \mathbf{d}_{P-1}^H]^H \quad (10)$$

$$\mathcal{C} := \underbrace{(\mathrm{diag}\{c_1, c_2, \ldots, c_{P-1}\}\mathbf{V})}_{=:\mathcal{V}} \otimes I_N \quad (11)$$

where $\otimes$ denotes the Kronecker product [7, p. 429]. Omitting the term $m = 0$ and using the definition of $\mathbf{d}_m$ from (4), it follows that

$$\mathcal{C}\mathcal{H} = \mathcal{D}. \quad (12)$$

In (8) $\mathbf{V}$ is a Vandermonde matrix with a rank of $L+1$ if $P-1 \ge L+1$ and α_i's are distinct [6, p. 274]. Since $c_m \ne 0 \, \forall m$, by [6, Result R4, p. 257], $\mathrm{rank}(\mathcal{V}) = \mathrm{rank}(\mathbf{V}) = L+1$. Finally, by [7, Property K6, p. 431], $\mathrm{rank}(\mathcal{C}) = \mathrm{rank}(\mathcal{V}) \times \mathrm{rank}(I_N) = N(L+1)$. Therefore, we can determine $\mathbf{h}(l)$'s uniquely. Define $\hat{\mathcal{D}}$ as in (12) with $\mathbf{d}_m$'s replaced with $\hat{\mathbf{d}}_m$'s. Then we have the channel estimate

$$\hat{\mathcal{H}} = (\mathcal{C}^H\mathcal{C})^{-1}\mathcal{C}^H\hat{\mathcal{D}}. \quad (13)$$

Precise knowledge of the channel length L is not required; an upperbound L_u suffices. Then we estimate $\mathbf{h}(i)$ for $0 \le i \le L_u$ with $\hat{\mathbf{h}}(i) \to 0$ m.s. for $i \ge L+1$ (= true channel length) as record length $T \to \infty$. Also, we do not need $c_m \ne 0$ for every m. We need at least $L+2$ nonzero c_m's. This can be accomplished by picking a "large" P and a suitable $\{c(n)\}$ (picked to satisfy a peak-to-average power constraint, e.g.,). Implicit in our approach (also in [1] and [2]) is the need at the receiver for synchronization with the transmitter's superimposed training sequence. •

A. Equalization

With $\hat{\mathbf{h}}(i)$ denoting the estimated $\mathbf{h}(i)$ and $\tilde{\mathbf{v}}(n) := \mathbf{v}(n) - \mathbf{m}$, define

$$\tilde{\mathbf{y}}(n) := \mathbf{y}(n) - \sum_{i=0}^{L} \hat{\mathbf{h}}(i)c(n-i) - \hat{\mathbf{m}} \approx \sum_{i=0}^{L} \mathbf{h}(i)s(n-i) + \tilde{\mathbf{v}}(n) \quad (14)$$

where $\hat{\mathbf{m}} := (1/T)\sum_{n=1}^{T}[\mathbf{y}(n) - \sum_{i=0}^{L} \hat{\mathbf{h}}(i)c(n-i)]$. That is, $\tilde{\mathbf{y}}(n)$ is obtained by removing the (estimated) contribution of the superimposed training and the dc-offset from the noisy data. Model (14) with the estimated channel is used to equalize the channel and to detect the information sequence. For the simulations of Section III we used a linear MMSE (minimum mean-square error) equalizer which also requires the knowledge of the correlation function of $\tilde{\mathbf{y}}(n)$. We estimate the noise variance σ_v^2 (see (**H2**)) as ($\mathrm{tr}\{A\}$ denotes trace of matrix A)

$$\hat{\sigma_v^2} := \mathrm{tr}\left\{\left[\frac{1}{NT}\sum_{n=1}^{T} \tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n)\right] - \sum_{i=0}^{L} \hat{\mathbf{h}}(i)\hat{\mathbf{h}}^H(i)\right\}. \quad (15)$$

(If (15) yields a negative result, we set it to zero.) The correlation function of $\tilde{\mathbf{y}}(n)$ can then be estimated using the estimated channel (instead of the less reliable sample averaging); only the zero lag correlation requires $\hat{\sigma_v^2}$.

III. Simulation Examples

A. Example 1

Consider a continuous-time channel $\tilde{h}(t)$ given by $\tilde{h}(t) = \sum_{i=1}^{4} a_i p_{4T_s}(t - \tau_i; 0.2)$ where T_s is the symbol interval, $p_{4T_s}(t; 0.2)$ denotes the raised-cosine pulse with roll-off factor 0.2 and length truncated to $4T_s$ (i.e., $p_{4T_s}(t; 0.2) = 0$ for $|t| > 2T_s$), the amplitudes a_i's are mutually independent, zero-mean, complex Gaussian with same variance for all i's, and delays τ_i's are mutually independent, uniformly distributed over $[0, 4T_s]$. The continuous-time channel $\tilde{h}(t)$ is sampled once every T_s seconds to yield the discrete-time channel $h(n) := \tilde{h}((n-1)T_s)$. Thus we have $N = 1$ in (1) leading to

$$y(n) = \sum_{l=0}^{7} h(l)[b(n-l) + c(n-l)] + v(n). \quad (16)$$

Let L_u be the upper bound on channel length $L = 7$. We take $L_u = 10$. The channel is randomly generated in each Monte Carlo. The input information sequence $\{b(n)\}$ is i.i.d. equiprobable 4-QAM (quadrature amplitude modulation) taking values $(\pm 1 \pm j)/\sqrt{2}$. The training sequence was chosen to have $P = 15$ with $c(n) = \sum_k \sqrt{\alpha}\delta(n - 15k)$ as in [2]; α is picked to yield a particular training-to-information sequence

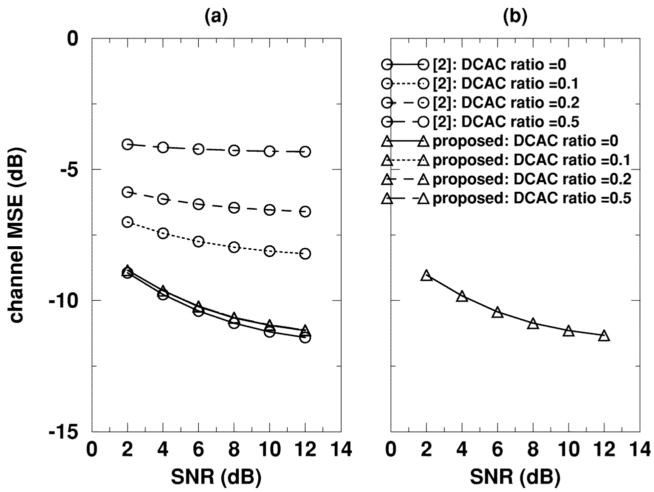

Fig. 1. (a) Example 1: Normalized channel MSE (17) based on $T = 150$ symbols per run, 100 Monte Carlo runs, $P = 15$, TIR $\alpha = 0.585$. $\mathrm{DCAC\ ratio} = [E\{v(n)\}]^2/(E\{|y(n) - v(n)|^2\})$. The curves for the proposed method for different DCAC ratios are overlaid (very close). (b) Normalized channel MSE for Example 2; the rest as for Fig. 1(a).

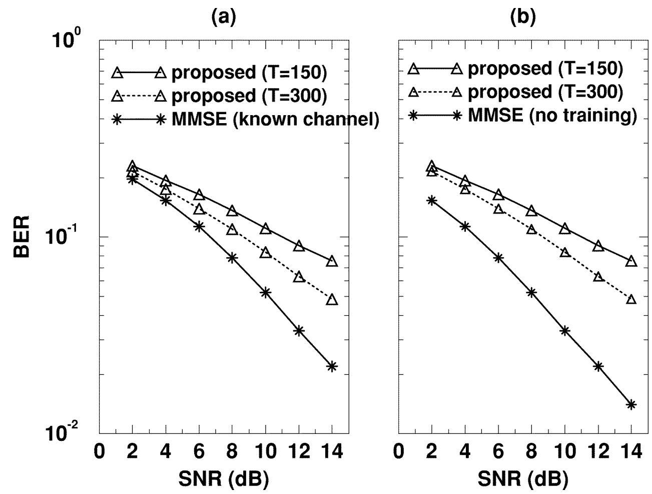

Fig. 2. Example 2. (a) Equalization performance using linear MMSE equalizers based on $T = 150$ or 300 symbols per run, 100 Monte Carlo runs, $P = 15$. $\mathrm{DCAC\ ratio} = 0$, TIR $\alpha = 0.585$. (b) Fig. 2(a) redrawn with the curve for the known-channel linear MMSE equalizer adjusted by 2 dB – no power is wasted in training.

power ratio (TIR) $\alpha = \sigma_c^2/\sigma_b^2$ where σ_b^2 and σ_c^2 denote the average power in the information sequence $\{b(n)\}$ and training sequence $\{c(n)\}$, respectively. Complex white zero-mean Gaussian noise was added to the received signal and scaled to achieve an SNR at the receiver (relative to the contribution of $\{s(n)\}$). A mean-value m was added to the noisy received signal to achieve a specified dc-offset to signal ac-component (DCAC) power ratio $m^2/(E\{|y(n) - v(n)|^2\})$. Normalized mean-square error in estimating the channel impulse response averaged over 100 Monte Carlo runs, was taken as the performance measure for channel estimation. It is defined as (before Monte Carlo averaging)

$$\mathrm{NCMSE} := \left[\sum_{l=0}^{L_u} \|h(l) - \hat{h}(l)\|^2\right] \left[\sum_{l=0}^{L_u} \|h(l)\|^2\right]^{-1}. \quad (17)$$

The simulation results are shown in Fig. 1(a) for various SNR's and DCAC power ratios for a record length of $T = 150$ symbols and a TIR of -2.33 dB ($\alpha = 0.585$). Our proposed method and that of [2] were simulated. It is seen that the proposed method is insensitive to the presence of the unknown mean m whereas the method of [2] is very sensitive. For $m = 0$, the performance of our method is slightly inferior to that of [2]. In the method of [2], $\hat{h}(l)$'s are estimated directly from data for $1 \le l \le L_u + 1 = 11$ whereas in our approach, we first estimate $\hat{\mathbf{d}}_m$'s for $1 \le m \le P - 1 = 14$ and then use (13). Since we estimate more variables (14 versus 11), this may account for the slightly inferior performance of our method for $m = 0$.

B. Example 2

This example is exactly as Example 1 except for the training sequence which was taken to be an m-sequence (maximal length pseudorandom binary sequence) of length 15 ($= P$), $c(n) = \sqrt{\alpha}\tilde{c}(n)$, $\{\tilde{c}(n)\}_{n=0}^{15} = \{-1,-1,-1,1,1,1,1,-1,1,-1,1,1,-1,-1,1\}$. The peak-to-average power ratio for this sequence is one (the best possible). The simulation results are shown in Fig. 1(b) for a record length of $T = 150$ symbols and a TIR of -2.33 dB ($\alpha = 0.585$). Only our proposed method was simulated since the method of [2] does not apply to this model. It is seen that as in Example 1, the proposed method is insensitive to the presence of the unknown mean m. Equalization performance (BER) of a linear MMSE equalizer based on the estimated channel (Example 2) is shown in Fig. 2(a) for two different record lengths of $T = 150$ and 300 symbols. The linear equalizer was designed as noted in Section II with equalizer length of 10 symbols and delay of five symbols. Also shown is the performance of a linear equalizer based upon perfect knowledge of the channel and noise variance. It is seen that the performance improves with record length. Note that for our choice of $\alpha = 0.585$, the SNR relative to $\{b(n)\}$ would be 2 dB less than the SNR shown in Fig. 2(a), which is relative to $\{s(n)\}$. To reflect this loss in SNR due to inclusion of the superimposed training, we redraw Fig. 2(a) as Fig. 2(b) with the SNR for the curve for the known-channel linear MMSE equalizer adjusted by 2 dB.

REFERENCES

[1] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Trans. Broadcasting*, vol. 46, pp. 170–176, June 2000.
[2] G. T. Zhou, M. Viberg, and T. McKelvey, "Superimposed periodic pilots for blind channel estimation," in *Proc. 35th Annu. Asilomar Conf. Signals Systems Computers*, Pacific Grove, CA, Nov. 5–7, 2001, pp. 653–657.
[3] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Mag.*, vol. 17, pp. 16–28, May 2000.
[4] A. V. Dandawate and G. B. Giannakis, "Asymptotic theory of mixed time average and kth-order cyclic-moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. IT-41, pp. 216–232, Jan. 1995.
[5] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[6] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
[7] B. Porat, *Digital Processing of Random Signals*. Upper Saddle River, NJ: Prentice-Hall, 1994.

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)

2. REPORT DATE **Oct. 6, 2003**

3. REPORT TYPE AND DATES COVERED
Peer Reviewed Reprint (Journal article)

4. TITLE AND SUBTITLE
On channel estimation using superimposed training and first-order statistics

5. FUNDING NUMBERS
DAAD19-01-1-0539

6. AUTHOR(S)
Jitendra K. Tugnait and Weilin Luo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Auburn University
Dept. of Electrical & Computer Engineering, 200 Broun Hall
Auburn, AL 36839

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
41703.14-CI

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Channel estimation for single-input multiple-output (SIMO) time-invariant channels is considered using only the first-order statistics of the data. A periodic (non-random) training sequence is added (superimposed) at a low power to the information sequence at the transmitter before modulation and transmission. Recently superimposed training has been used for channel estimation assuming no mean-value uncertainty at the receiver and using periodically inserted pilot symbols. We propose a different method that allows more general training sequences and explicitly exploits the underlying cyclostationary nature of the periodic training sequences. We also allow mean-value uncertainty at the receiver. Illustrative computer simulation examples are presented.
This paper has been published in the *IEEE Communications Letters*, vol. CL-7, pp. 413-415, Sept. 2003.

14. SUBJECT TERMS
Channel estimation, superimposed training.

15. NUMBER OF PAGES
4

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	**UL**

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102